MODELOS DE PROPORÇÕES COM SUPERDISPERSÃO E EXCESSO DE ZEROS

Adriano Ferreti Borgatto, Prof. Dr. do Departamento de Informática e Estatística, INE–UFSC, e-mail: borgatto@inf.ufsc.br.

RESUMO: Inúmeras são as áreas em que se obtêm dados de proporções, como por exemplo Entomologia. Este trabalho consiste em analisar um conjunto de dados obtido de um ensaio de controle biológico, em que o interesse está na quantidade de ovos parasitados dentre um total de m_i ovos. Assumindo que Y_i é uma variável aleatória "número de ovos parasitados" em m_i ovos, com observações (y_i, m_i) e $p_i = \frac{y_i}{m_i}$, representando a proporção de ovos parasitados, i = 1, 2,..., n, sendo n o número total de unidades experimentais. Assume-se inicialmente que a variável aleatória Y_i tem distribuição binomial e associando a covariável ao modelo através da função logística, o modelo não se ajusta bem devido à superdispersão gerada pela variabilidade da média. Com isso, uma forma de modelar essa variabilidade é alterar o modelo probabilístico. Uma possibilidade para ajustar os dados com superdispersão é assumir uma distribuição de probabilidade sobre o parâmetro p_i da distribuição binomial (Hinde & Demétrio, 1998a). Assumindo que $Y_i \mid P \sim \mathrm{Bin}(m_i, P)$ e $P \sim \mathrm{Beta}(a, b)$, $a > 0$ e $b > 0$, assim incondicionalmente Y_i tem distribuição beta-binomial. Usando a reparametrização (Williams, 1982) $p_i = \frac{a}{a+b}$ e $\delta = \frac{1}{a+b}$, o modelo beta-binomial fica com média dada por $E(Y_i) = m_i p_i$ e variância $\mathrm{Var}(Y_i) = m_i p_i (1 - p_i)\left(\frac{m_i \delta + 1}{\delta + 1}\right)$. Além da superdispsersão gerada pela variabilidade da média o conjunto de dados apresenta excesso de zeros. Uma abordagem para dados de proporções com superdispersão e excesso de zeros, é pensar sobre a incapacidade natural de algumas fêmeas em parasitar os ovos do hospedeiro. Nesse caso, a variável aleatória Y_i pode ser representada por $Y_i = V_i(1 - H_i)$ (Ghosh et al., 2005), sendo H_i uma variável aleatória Bernoulli, isto é, $H_i \sim \mathrm{Ber}(w_i)$, em que w_i é a probabilidade de uma fêmea parasitar e V_i uma variável aleatória com distribuição beta-binomial. Tem-se, então, que Y_i assume uma distribuição beta-binomial inflacionada de zeros (ZIBB).